ISBN 82-553-0508-4

No 2
1983

DEMOGRAPHIC INCIDENCE RATES AND ESTIMATION OF INTENSITIES WITH INCOMPLETE INFORMATION

by

Ørnulf Borgan
Institute of Mathematics
University of Oslo

and

Henrik Ramlau-Hansen
Laboratory of Actuarial Mathematics
University of Copenhagen

Summary

Many population processes in demography, epidemiology and other fields can be represented by a time-continuous Markov chain model with a finite state space. If we have complete information on the life history of a cohort, the intensities of the Markov model may be estimated by the occurrence/exposure rates or by nonparametric techniques. In many situations we have only incomplete information, however, since the occurrences and the total exposure are known, but not the distribution of the latter over the various separate statuses. Methods to handle such problems, e.g. the demographic incidence rates, are known in the literature. Their statistical properties are only vaguely known, however. The present paper gives a thorough presentation of the theory of these methods, and provide rigorous proofs of their statistical properties using stochastic process theory.

AMS 1980 subject classifications. 62M05, 62P99

Key words and phrases. Asymptotic theory, cohort analysis, counting processes, cumulative incidence rates, martingales, occurrence/exposure rates, time-continuous Markov chains.

1. Introduction

In demography and fields with a similar methodology, one can frequently represent a population process on the individual level by a time-continuous Markov chain with a finite state space. The states of the Markov chain represent the demographic statuses and the jumps between the states correspond to the demographic events. The time parameter of the Markov process may be the age of an individual, the time since a specific event, or some such quantity. We will call it <u>seniority</u>.

Given this general framework, the population phenomena may be described by the transition intensities of the model. Consequently, it is of great interest to estimate these functions. In principle, this is an easy task if the individuals under study are watched continuously over some period of time. One may then use the classical methods based on occurrence/exposure rates (for a review, see Hoem, 1976), or for small populations nonparametric techniques developed by Aalen (1978) and others. Quite often, however, one does not obtain the detailed data required to use any of these methods. For instance, consider a model for the first marriage in a female birth cohort. To compute occurrence/exposure rates for the first marriage intensity, one needs information on the first marriages by age of the bride <u>and</u> data on the distribution of the cohort over the marital statuses at each age. In such situations, quite often one has sufficient information about the number of occurrences of the events (marriages), but the numbers of person-years lived in the various statuses, i.e. the exposures, frequently are unknown, however, and some other method must be sought to analyse the data.

To handle such problems, demographers have for a long time computed incidence rates, which are calculated as the number of occurrences of a specific event during a given period divided by the mean population alive during that period, for all statuses specified, taken together. Those not "at risk" for the event in question are not excluded from the population in the denominator, as is the case for occurrence/exposure rates. When the data are from a closed cohort, such an approach is well motivated, since then the cumulative incidence rate often represents the prevalence of the event studied (Hoem, 1978).

For simple situations, like the first marriage model mentioned briefly above, it has been known for some time how it is possible to use the cumulative incidence rates to compute estimates for the intensities themselves when the mortality is non-differential. Recently, Finnäs (1980) has shown how this method can be generalized to any Markov chain model with a single absorbing state when everyone has the same status at the outset, i.e. at seniority 0. Finnäs (1980) also showed somewhat informally that the cumulative incidence rates, as well as the estimators for the intensities themselves, are consistent for what they are said to estimate, and that they are asymptotically normally distributed. He gave only very indirect expressions for the asymptotic variances, however, and therefore he was not able to derive estimators for these.

The main purpose of the present paper is to give a thorough presentation of the theory of incidence rates, and to use stochastic process theory to provide rigorous proofs for Finnäs' (1980) results and some generalizations of these. We are also able to derive explicit expressions for the asymptotic variances and to provide estimators for these.

The plan of the paper is as follows. In the next section we introduce our Markov chain model, which is slightly more general than the one considered by Finnäs (1980). We describe in Section 3 how the classical incidence rates and Finnäs' (1980) estimators for the intensities appear in our set-up. In Section 4 we show how the situation at hand may be formulated in a counting process framework, and review some useful results which emerge from this formulation. Our version of the cumulative incidence rate is introduced in Section 5, where we also discuss its statistical properties. In the Sections 6 and 7 we discuss nonparametric estimation of the integrated intensities and estimation of piecewise constant intensities. A numerical example for the first marriage model is given in Section 8. Three appendices contain some of the more technical derivations.

We will use results from the theory of counting processes, martingales and stochastic integrals without further comment. The reviews given by Aalen (1978), Aalen and Johansen (1978) and Gill (1980) should cover our needs. An approach to the theory of counting processes attempting to minimize the dependence on general martingale theory is given by Jacobsen (1982).

2. The model

Assume that we study a closed cohort of n individuals, and that the phenomena of interest can be described by a time-continuous Markov chain model with time domain $[0,z]$ and with a finite state space $\underline{\underline{I}}$. The transition probabilities are denoted $P_{ij}(x,y)$, and the corresponding intensities, or forces of transition, are $\alpha_{ij}(x)=\lim_{y\downarrow x} P_{ij}(x,y)/(y-x)$ for $i\neq j$. We make the necessary assumptions which ensure that the Markov chain is well behaved, e.g., we assume that $P_{ij}(x,x)=\delta_{ij}$, where δ_{ij} is a Kronecker delta, and that the intensities are left-continuous functions with right-hand limits on $[0,z]$. This also makes $\int_0^z\alpha_{ij}(s)ds<\infty$ for all $i\neq j$.

Assume that $\underline{\underline{I}}$ contains an absorbing subset $\underline{\underline{R}}$ of states, i.e. $\sum_{k\in\underline{\underline{K}}} P_{ik}(x,y)=0$ for $0\leqslant x\leqslant y\leqslant z$ for all $i\in\underline{\underline{R}}$ where $\underline{\underline{K}}=\underline{\underline{I}}\setminus\underline{\underline{R}}$. Finnäs (1980) considered the situation where the absorbing subset $\underline{\underline{R}}$ consists of a single state called "dead". We also assume that $\sum_{j\in\underline{\underline{R}}} \alpha_{ij}(\cdot)=\mu(\cdot)$ for all $i\in\underline{\underline{K}}$, so that there is <u>non-differential</u> risk of transition from $\underline{\underline{K}}$ to $\underline{\underline{R}}$. Hoem (1969) has proved that this assumption implies that

$$P_{ij}(x,y) = \bar{P}_{ij}(x,y)p(x,y) \quad \text{for} \quad i,j\in\underline{\underline{K}}, \tag{2.1}$$

where $p(x,y)=\exp(-\int_x^y\mu(s)ds)$ and $\bar{P}_{ij}(x,y)$ are the transition probabilities of the <u>partial</u> Markov chain with state space $\underline{\underline{K}}$ obtained by substituting 0 for α_{ij} for all (i,j) with $j\in\underline{\underline{R}}$. Note that in the present situation $\bar{P}_{ij}(x,y)$ is also the conditional probability that an individual in state i at seniority x is in state j at seniority y, given that he is still in $\underline{\underline{K}}$ at the latter seniority.

Let us assume that the individuals start in a state in $\underline{\underline{K}}$ at seniority 0 independently of each other and according to an initial distribution $(\pi_k;k\in\underline{\underline{K}})$. (Finnäs, 1980, considered the case where

$\pi_k=\delta_{1k}$ for some state $1\in\underline{\underline{K}}$.) Define $\bar{P}_i(x) = \sum_{k\in\underline{\underline{K}}} \pi_k \bar{P}_{ki}(0,x)$ and let $P_i(x) = \bar{P}_i(x)p(x)$ for $i\in\underline{\underline{K}}$, where $p(x)=p(0,x)$.

Let $\beta_{ij}(x)$ be the expected number of transitions directly from a state $i\in\underline{\underline{K}}$ to another state $j\in\underline{\underline{K}}$ in the seniority interval $[0,x]$ for an individual who is still in $\underline{\underline{K}}$ at age x. Then

$$\beta_{ij}(x) = \int_0^x \alpha_{ij}(s)\bar{P}_i(s)ds , \tag{2.2}$$

and using the Kolmogorov differential equations one gets

$$\bar{P}_i(x) = \beta_{\cdot i}(x) - \beta_{i\cdot}(x) + \pi_i , \tag{2.3}$$

as in Finnäs(1980), where the dots here and in what follows signify summation over all $k\in\underline{\underline{K}}\setminus\{i\}$, when not another definition is explicitely given.

Assume that out of the n individuals, N_k start out in state $k\in\underline{\underline{K}}$, so that

$$N_k/n \xrightarrow{P} \pi_k \quad \text{as} \quad n \to \infty .$$

3. Estimation, assuming piecewise constant intensities

In demography and other fields it is common to adopt the "semi-parametric" assumption of piecewise constant intensities (e.g. Hoem, 1976; Hoem & Jensen, 1982).

Let therefore in this section $0=a_1<a_2<\ldots<a_{R+1}=z$ be a partitioning of the seniority interval $[0,z]$ into subintervals $(a_r,a_{r+1}]$; $r=1,2,\ldots,R$; and assume that the intensities are constant on each of the subintervals, i.e. $\alpha_{ij}(x)=\alpha_{ijr}$ for $x\in(a_r, a_{r+1}]$. Let $F_{ij}(a_r)$ be the number of transitions direct from state i to state j in $\underline{\underline{K}}$ during the r-th subinterval, and let $L_i(a_r)$ be the total number of person-years lived in state i during the same interval. We write $L_{\cdot}(a_r)=\sum_{i\in\underline{\underline{K}}} L_i(a_r)$. Then the

($\underline{\underline{K}}$-restricted) incidence rate is

$$b_{ij}(a_r) = F_{ij}(a_r)/L_{\bullet}(a_r) \tag{3.1}$$

for $i,j\in\underline{\underline{K}}$, $i\neq j$, and the ($\underline{\underline{K}}$-restricted) cumulative incidence rate up to seniority $x=a_{m+1}$ is given by

$$\bar{B}_{ij}(x) = \sum_{r=1}^{m} b_{ij}(a_r)(a_{r+1}-a_r). \tag{3.2}$$

As noted by Hoem (1978), this is an estimator for $\beta_{ij}(x)$, given by (2.2).

Combining (2.2) and (2.3), we find that

$$\bar{\alpha}_{ijm} = \frac{b_{ij}(a_m)}{N_i/n+\bar{B}_{\bullet i}(a_m)-\bar{B}_{i\bullet}(a_m)+[b_{\bullet i}(a_m)-b_{i\bullet}(a_m)](a_{m+1}-a_m)/2} \tag{3.3}$$

is an estimator for α_{ijm}, $i,j\in\underline{\underline{K}}$, $i\neq j$. The argument is similar to the one given by Finnäs (1980) and is omitted.

Note that the estimators given by (3.1) through (3.3), as well as $\beta_{ij}(x)$ defined in (2.2), depend on the subdivision of the state space into the two subsets $\underline{\underline{K}}$ and $\underline{\underline{R}}$. Thus if it is possible to split up the state space into two other components $\underline{\underline{K}}'$ and $\underline{\underline{R}}'$, where $\underline{\underline{R}}'$ is an absorbing subset of states and where the risk of transition from $\underline{\underline{K}}'$ to $\underline{\underline{R}}'$ is nondifferential, it is possible to get new estimators for α_{ijm} for $i,j\in\underline{\underline{K}}\cap\underline{\underline{K}}'$. Of course the different estimators have different data requirements and different properties (cf. Section 7).

4. A counting process formulation

Denote by $K_{ij}(x)$ the number of transitions directly from state i to state j experienced by the cohort in the seniority interval $[0,x]$, and let $Y_i(x)$ be the number of individuals in state i

"just before" seniority x, i.e. $Y_i(\cdot)$ is left-continuous. Moreover, let F_x be the σ-algebra generated by $(N_k; k \in \underline{\underline{K}})$ and $(K_{ij}(s);$ $0 \leq s \leq x,\ i,j \in \underline{\underline{I}},\ i \neq j)$. Then $(K_{ij}(x);\ 0 \leq x \leq z,\ i,j \in \underline{\underline{I}},\ i \neq j)$ is a multivariate counting process where $K_{ij}(\cdot)$ has the intensity process $\alpha_{ij}(\cdot)Y_i(\cdot)$ relative to the increasing family of σ-algebras (F_x). By the theory of counting processes this implies that $(M_{ij}(x);$ $0 \leq x \leq z,\ i,j \in \underline{\underline{I}},\ i \neq j)$, given by

$$M_{ij}(x) = K_{ij}(x) - \int_0^x \alpha_{ij}(s)Y_i(s)ds, \tag{4.1}$$

are orthogonal square integrable martingales with respect to (F_x). The variance process $\langle M_{ij}\rangle(\cdot)$ of M_{ij} is

$$\langle M_{ij}\rangle(x) = \int_0^x \alpha_{ij}(s)Y_i(s)ds, \tag{4.2}$$

which means that $M_{ij}^2 - \langle M_{ij}\rangle$ is a square integrable martingale.

For the situation where the information structure is (F_x), i.e., where the $\{K_{ij}\}$ as well as the $\{Y_i\}$ are observable, Aalen (1978) approached the estimation problem as follows. Since by (4.1) we may write symbolically

$$\text{"}dK_{ij}(x) = \alpha_{ij}(x)Y_i(x)dx + \text{noise"},$$

a natural estimator for the cumulative intensity $\gamma_{ij}(x) = \int_0^x \alpha_{ij}(s)ds$ is $\int_0^x [Y_i(s)]^{-1}dK_{ij}(s)$. However, one may have $Y_i(x)=0$ for some i and x. Therefore Aalen introduced the indicator processes $J_i(x) = I(Y_i(x)>0)$, and defined the estimators formally by

$$\hat{\gamma}_{ij}(x) = \int_0^x J_i(s)[Y_i(s)]^{-1}dK_{ij}(s), \tag{4.3}$$

where $0/0$ is interpreted as 0. Results on uniform consistency and asymptotic normality of these estimators can be found in Aalen (1978).

In the situation considered in this paper, one does not observe the $\{Y_i\}$ and all the $\{K_{ij}\}$, however, so we are not able to calculate the Nelson-Aalen estimators (4.3). Our observational plan is to observe (N_k), the K_{ij} with $i,j \in \underline{\underline{K}}$, and $A=\sum_{i\in\underline{\underline{K}}}\sum_{j\in\underline{\underline{R}}} K_{ij}$, the process counting the number of absorptions in $\underline{\underline{R}}$. This means that the individual $\{Y_i; i\in\underline{\underline{K}}\}$ are not observed, but $Y_{\bullet}(x)=\sum_{i\in\underline{\underline{K}}} Y_i(x)$ $=n-A(x-)$ is still observable. Hence, the observed "history" of the cohort in the seniority interval $[0,x]$ can be described by the σ-algebra $\mathcal{G}_x$ generated by $(N_k; k\in\underline{\underline{K}})$ and $(A(s), K_{ij}(s);$ $0\leqslant s\leqslant x,\ i,j\in\underline{\underline{K}},\ i\neq j)$.

5. Nonparametric estimation of β_{ij}

Within the counting process framework of Section 4, we now consider nonparametric estimation of the $\{\beta_{ij}\}$ defined by (2.2). Introduce

$$B_{ij}(x) = \int_0^x J(s)[Y_{\bullet}(s)]^{-1} dK_{ij}(s), \tag{5.1}$$

where $J(x)=I(Y_{\bullet}(x)>0)$, as an estimator for $\beta_{ij}(x)$ for $i,j\in\underline{\underline{K}}$, $i\neq j$. Note that (5.1) can be seen as the limit of the cumulative incidence rate in (3.2) as the number of subintervals goes to infinity and their lengths to zero. The real justification for using B_{ij} as an estimator for β_{ij} is given in the following proposition and theorems, however.

Proposition 5.1.

$EB_{ij}(x)=\int_0^x \bar{P}_i(s)\alpha_{ij}(s)P(Y_{\bullet}(s)>0)ds=\int_0^x P(Y_{\bullet}(s)>0)d\beta_{ij}(s)$ for $i,j\in\underline{\underline{K}}$, $i\neq j$, $x\in[0,z]$, where $P(Y_{\bullet}(x)>0)=1-(1-p(x))^n$.

Proof: By (4.1)

$$B_{ij}(x) = \int_0^x \alpha_{ij}(s)Y_i(s)[Y_{\bullet}(s)]^{-1}ds+\int_0^x J(s)[Y_{\bullet}(s)]^{-1}dM_{ij}(s). \tag{5.2}$$

Here, the final term is a stochastic integral with respect to a square integrable martingale, and hence itself a zero mean martingale. Thus $EB_{ij}(x)=\int_0^x \alpha_{ij}(s)E[Y_i(s)/Y_\bullet(s)]ds$. Since we have assumed a non-differential risk of transition from $\underline{\underline{K}}$ to $\underline{\underline{R}}$, $(Y_i(s);\ i\in\underline{\underline{K}})$ is multinomially distributed with parameters $(\bar{P}_i(s);i\in\underline{\underline{K}})$, conditionally on $Y_\bullet(s)>0$. Therefore, $EB_{ij}(x) = \int_0^x\alpha_{ij}(s)\bar{P}_i(s)P(Y_\bullet(s)>0)ds$, where $P(Y_\bullet(x)>0)=1-(1-p(x))^n$. □

To discuss asymptotic properties of (5.1), consider the sequence of counting processes we get by letting $n\to\infty$, and index all relevant quantities by n. Then the next result shows that $B_{ij}^{(n)}$ is an uniformly consistent estimator for β_{ij}.

<u>Theorem 5.2</u>. Let $B_{ij}^{(n)}(x)$ and $\beta_{ij}(x)$ be given by (5.1) and (2.2), respectively. Then under the assumptions of Section 2

$$\sup_{x\in[0,z]}|B_{ij}^{(n)}(x) - \beta_{ij}(x)| \overset{P}{\to} 0$$

as $n \to \infty$.

<u>Proof</u>: By (5.2) it is sufficient to prove that

$$\sup_{s\in[0,z]}|Y_i^{(n)}(s)[Y_\bullet^{(n)}(s)]^{-1}- \bar{P}_i(s)| \overset{P}{\to} 0 \tag{5.3}$$

and

$$\sup_{x\in[0,z]}|\int_0^x J^{(n)}(s)[Y_\bullet^{(n)}(s)]^{-1}dM_{ij}^{(n)}(s)| \overset{P}{\to} 0 \tag{5.4}$$

as $n\to\infty$. By standard results (5.3) is fulfilled. Using Lenglart's (1977) inequality (cf. Andersen and Gill, 1982, Appendix I) and (4.2), we have

$$P\{\sup_{x\in[0,z]}|\int_0^x J^{(n)}(s)[Y_\bullet^{(n)}(s)]^{-1}dM_{ij}^{(n)}(s)|^2>\varepsilon\}$$

$$< \frac{\delta}{\varepsilon} + P\{\int_0^z J^{(n)}(s)[Y_{\cdot}^{(n)}(s)]^{-2} d\langle M_{ij}^{(n)}\rangle(s) > \delta\}$$

$$= \frac{\delta}{\varepsilon} + P\{\int_0^z \alpha_{ij}(s) Y_i^{(n)}(s)[Y_{\cdot}^{(n)}(s)]^{-2} ds > \delta\}$$

for all $\varepsilon, \delta > 0$. Since $\int_0^z \alpha_{ij}(s) Y_i^{(n)}(s)[Y_{\cdot}^{(n)}(s)]^{-2} ds \xrightarrow{P} 0$ as $n \to \infty$, (5.4) is also fulfilled. □

Theorem 5.2 is a much stronger result than the one proved by Finnäs (1980), who showed that $\bar{B}_{ij}$ given by (3.2) is pointwise consistent of a Riemann sum corresponding to the integral (2.2) in the piecewise constant intensity set-up of Section 3.

Let us then turn to the problem of proving an asymptotic distributional result for the cumulative incidence rates (5.1). The usual way of proving such results within the counting process framework, is to apply some version of the martingale central limit theorem (Rebolledo, 1978, 1980). However, in our case it seems difficult to proceed in this way, and we will therefore apply a Skorohod construction as in Breslow and Crowley (1974, Theorem 4). For this purpose, we need the following central limit theorem for the asymptotic distribution of the number of transitions between the various states in a general time-continuous Markov chain.

The weak convergence takes place in the space $D^m[0,z]$ of m-dimensional functions on $[0,z]$ with right-continuous real valued components with left-hand limits equipped with the Skorohod product topology (see Billingsley, 1968), where m is the number of component processes.

Theorem 5.3. Consider n independent copies of the same time continuous Markov chain on $[0,z]$ with a finite state space $\underline{\underline{I}}$ and with transition intensities and probabilities $\alpha_{ij}(s)$ and $P_{ij}(s,t)$, respectively. Assume that there exists a constant $C>0$

such that $\alpha_{ij}(s) \leq C$ for $s \in [0,z]$ and for all $i,j \in \underline{\underline{I}}$, $i \neq j$. Let $K_{ij}^{(n)}(t)$ denote the total number of transitions from state i to state j in $[0,t]$ and introduce the normalized number of transitions

$$X_{ij}^{(n)}(t) = \sqrt{n}\{n^{-1}K_{ij}^{(n)}(t) - \nu_{ij}(t)\},$$

where $\nu_{ij}(t) = \int_0^t P_i(\sigma)\alpha_{ij}(\sigma)d\sigma$ with $P_i(s) = \sum_{j \in \underline{\underline{I}}} \pi_j P_{ji}(0,s)$ and $\{\pi_i;\ i \in \underline{\underline{I}}\}$ the initial distribution. Then $X^{(n)} = (X_{ij}^{(n)};\ i,j \in \underline{\underline{I}},\ i \neq j)$ converges weakly to a mean zero Gaussian process $X = (X_{ij};\ i,j \in \underline{\underline{I}},\ i \neq j)$ with covariance structure given for $s \leq t$ by

$$\begin{aligned} \mathrm{Cov}(X_{ij}(s), X_{kl}(t)) &= \int_0^s \int_0^\tau P_{jk}(\sigma,\tau)\alpha_{kl}(\tau)d\nu_{ij}(\sigma)d\tau \\ &+ \int_s^t \int_0^s P_{jk}(\sigma,\tau)\alpha_{kl}(\tau)d\nu_{ij}(\sigma)d\tau \\ &+ \int_0^s \int_0^\sigma P_{li}(\tau,\sigma)\alpha_{ij}(\sigma)d\nu_{kl}(\tau)d\sigma \qquad (5.6) \\ &+ \delta_{ik}\delta_{jl}\nu_{ij}(s) - \nu_{ij}(s)\nu_{kl}(t). \end{aligned}$$

<u>Remark</u>: Note that no assumptions of non-differential risk of transition are necessary for this theorem to hold true. Moreover, the condition that the intensities are uniformly bounded is automatically fulfilled under the assumptions made on the intensities in Section 2. $\square$

We feel that this theorem must have been proved before, but we have not been able to find a proof in the literature. Therefore we provide one in Appendix A.

Let us then apply this theorem to prove weak convergence of the cumulative incidence rates given in (5.1). Let $X_{ij}^{(n)}$ for $i,j \in \underline{\underline{K}}$, $i \neq j$, be as in Theorem 5.3 and introduce

$$U^{(n)}(x) = \sqrt{n}\{n^{-1}Y_{\bullet}^{(n)}(x) - p(x)\}\ . \qquad (5.7)$$

The corresponding limiting process is denoted U. Then by Theorem

5.3 it is easily seen that $(U^{(n)}, X_{ij}^{(n)};\ i,j\in\underline{\underline{K}},\ i\neq j)$ converges weakly to $(U, X_{ij};\ i,j\in\underline{\underline{K}},\ i\neq j)$, and that

$$\mathrm{Cov}(X_{ij}(x),U(y)) = \begin{cases} \int_0^x p(s,y)d\nu_{ij}(s)-p(y)\nu_{ij}(x) \text{ for } x\leq y & (5.8)\\ \nu_{ij}(x)-\nu_{ij}(y)+\int_0^y p(s,y)d\nu_{ij}(s)-p(y)\nu_{ij}(x) \text{ for } y<x \end{cases}$$

where $\nu_{ij}(x)=\int_0^x P_i(s)\alpha_{ij}(s)ds$, and for $x\leq y$

$$\mathrm{Cov}(U(x),U(y)) = p(y)(1-p(x)). \tag{5.9}$$

Note that by (5.8), $X_{ij}(x)$ and $U(y)$ are positively correlated when $x\leq y$ as one would expect.

Also note that under the assumptions of Section 2, $p(z)>0$ and $\nu_{ij}(z)<\infty$ for all $i,j\in\underline{\underline{K}}$. We are then able to state the following result.

<u>Theorem 5.4</u>. Under the assumptions of Section 2, the multivariate process $(\sqrt{n}(B_{ij}^{(n)}-\beta_{ij});\ i,j\in\underline{\underline{K}},\ i\neq j)$ converges weakly to a zero mean Gaussian process $(Z_{ij};\ i,j\in\underline{\underline{K}},\ i\neq j)$, where

$$\begin{aligned} Z_{ij}(x) &= -\int_0^x U(s)[p(s)]^{-2}d\nu_{ij}(s) + X_{ij}(x)[p(x)]^{-1} \\ &\quad -\int_0^x X_{ij}(s)\mu(s)[p(s)]^{-1}ds. \end{aligned} \tag{5.10}$$

For $x\leq y$, the covariance structure of the limiting process is given by

$$\begin{aligned} \mathrm{Cov}(Z_{ij}(x),Z_{kl}(y)) =& \int_0^x\int_0^u P_{jk}(r,u)\alpha_{kl}(u)[p(r)p(u)]^{-1}d\nu_{ij}(r)du \\ &+\int_0^x\int_0^r P_{li}(u,r)\alpha_{ij}(r)[p(r)p(u)]^{-1}d\nu_{kl}(u)dr \qquad (5.11)\\ &+\int_x^y\int_0^x P_{jk}(r,u)\alpha_{kl}(u)[p(r)p(u)]^{-1}d\nu_{ij}(r)du \\ +\delta_{ik}\delta_{jl}\int_0^x[p(u)]^{-2}d\nu_{ij}(u) &-\int_0^x\int_0^u[p(r)]^{-2}[p(u)]^{-1}d\nu_{ij}(r)d\nu_{kl}(u) \\ &-\int_0^x\int_0^u [p(r)]^{-2}[p(u)]^{-1}d\nu_{kl}(r)d\nu_{ij}(u) \\ &-\left(\int_0^x[p(r)]^{-2}d\nu_{ij}(r)\right)\left(\int_x^y[p(u)]^{-1}d\nu_{kl}(u)\right). \end{aligned}$$

<u>Proof</u>: The convergence of $(\sqrt{n}(B_{ij}^{(n)}-\beta_{ij}))$ to (Z_{ij}), given by (5.10) follows by a Skorohod construction just as in the proof of Theorem 4 in Breslow and Crowley (1974). (Be aware of the misprint pointed out by Gill, 1981, p.4.) The covariance structure (5.11) follows by straightforward computations given in some detail in Appendix B. □

Note that by (2.1) and (2.2), $d\nu_{ij}(u)=p(u)d\beta_{ij}(u)$, so the covariance structure (5.11) may alternatively be expressed in terms of integration with respect to the β_{ij}. In general, we will need data on the individual level to estimate the covariances given by (5.11). However, in the case where only one transition from i to j is possible for each individual, one finds for $x \leqslant y$

$$\begin{aligned}\mathrm{Cov}(Z_{ij}(x),Z_{ij}(y)) = &\int_0^x [p(u)]^{-1}d\beta_{ij}(u) - 2\int_0^x\int_0^u [p(r)]^{-1}d\beta_{ij}(r)d\beta_{ij}(u)\\ &- \left(\int_0^x [p(r)]^{-1}d\beta_{ij}(r)\right)\left(\int_x^y d\beta_{ij}(u)\right),\end{aligned}$$

which may be estimated by

$$\begin{aligned}&n\int_0^x J^{(n)}(u)[Y^{(n)}(u)]^{-2}dK_{ij}^{(n)}(u)\\ &- 2n\int_0^x\int_0^u J^{(n)}(r)J^{(n)}(u)[Y^{(n)}(r)]^{-2}[Y^{(n)}(u)]^{-1}dK_{ij}^{(n)}(r)dK_{ij}^{(n)}(u)\\ &-n\left(\int_0^x J^{(n)}(r)[Y^{(n)}(r)]^{-2}dK_{ij}^{(n)}(r)\right)\left(\int_x^y J^{(n)}(u)[Y^{(n)}(u)]^{-1}dK_{ij}^{(n)}(u)\right).\end{aligned}$$

6. Nonparametric estimation of the integrated intensities

In this section, we study nonparametric estimation of the integrated intensities $\gamma_{ij}(x)=\int_0^x\alpha_{ij}(s)ds$. A natural first approach to this (which also was the first one we tried) is to formulate the situation in terms of the counting process set-up of Section 4, and to see whether some general results from the theory of counting processes may be applied. By the last remark in Section 4, $(A, K_{ij};$ $i,j\in\underline{\underline{K}},\ i\neq j)$ is a multivariate counting process relative to the increasing family of σ-algebras (G_x). By the inovation theorem (e.g. Aalen, 1978, Theorem 3.4), it follows that $K_{ij}(x)$ has the intensity process $\alpha_{ij}(x)E(Y_i(x)|G_x)$, and $A(x)$ has the intensity process $\mu(x)Y_{\bullet}(x)$. It would have been nice if $E(Y_i(x)|G_x)$ only depended on $(N_k; k\in\underline{\underline{K}})$ and $\{A(s),\ K_{ij}(s);\ i,j\in\underline{\underline{K}},\ i\neq j;\ 0\leq s<x\}$. For if this were the case, then the multiplicative structure leading to (4.3) would be retained, and we could derive estimators for the $\gamma_{ij}(\cdot)$ and study their properties by the methods of Aalen (1978). Unfortunately this is not the case, however, as is seen by the following argument:

Let $0\leq\tau_1<\tau_2<\ldots$ be the successive seniorities for the transitions in the multivariate counting process $(A, K_{ij};\ i,j\in\underline{\underline{K}},$ $i\neq j)$, and consider the situation where $\tau_1\leq x<\tau_2$, i.e. , exactly one transition has occurred up to senority x. Assume that this was an absorbtion into $\underline{\underline{R}}$, so that $\Delta A(\tau_1)=1$. Introduce $A_i=\sum_{j\in\underline{\underline{R}}}K_{ij}$. Then

$$E(Y_i(x)|G_x)=N_i-E(\Delta A_i(\tau_1)|G_x).$$

Now $P(\Delta A_i(\tau_1)=1|G_x)=P(\Delta A_i(\tau_1)=1,\ \tau_2>x|N_k,\ k\in\underline{\underline{K}};\ \tau_1;\ \ \Delta A(\tau_1)=1)/$ $P(\tau_2>x|N_k,\ k\in\underline{\underline{K}};\ \tau_1;\ \ \Delta A(\tau_1)=1)$. If $\psi_i=\sum_{j\in\underline{\underline{K}}\setminus\{i\}}\alpha_{ij}+\mu$, we find that the numerator equals $(N_i/n)\exp(-\int_{\tau_1}^{x}\psi_i(s)(N_i-1)ds)\times$ $\exp(-\sum_{j\in\underline{\underline{K}}\setminus\{i\}}\int_{\tau_1}^{x}\psi_j(s)N_j ds)$, and therefore

$$E(Y_i(x)|G_x)=N_i-N_i\exp(\int_{\tau_1}^{x}\psi_i(s)ds)/[\sum_{j\in\underline{\underline{K}}} N_j\exp(\int_{\tau_1}^{x}\psi_j(s)ds].$$

Thus even for this simple situation, we get a rather complicated expression for $E(Y_i(x)|G_x)$. It seems that the inovation theorem is of little use for our purpose, and we will derive our estimators from more intuitive arguments.

By (2.3), a natural "estimator" for the unobserved number $Y_i(x)$ at risk is

$$\tilde{Y}_i(x) = Y_{\bullet}(x)[B_{\bullet i}(x-) - B_{i\bullet}(x-) + N_i/n]. \tag{6.1}$$

Therefore, in analogy with (4.3), we propose

$$\tilde{\gamma}_{ij}(x) = \int_0^x \tilde{J}_i(s)[\tilde{Y}_i(s)]^{-1}dK_{ij}(s) \tag{6.2}$$

as an estimator for $\gamma_{ij}(x)$, where $\tilde{J}_i(x)=I(\tilde{Y}_i(x)>0)$. One should realize that the "estimated number at risk" $\tilde{Y}_i(x)$ can be negative. Consider for example the first marriage model, where $\underline{\underline{K}}=\{0,1\}$, $\underline{\underline{R}}=\{2\}$, and no transition from state 1 to state 0 is possible. Let $N_0=N_1=1$, and assume that at τ_1 we have a transition $1\to 2$, at $\tau_2>\tau_1$ a transition $0\to 1$, and at $\tau_3>\tau_2$ a transition $1\to 2$. Then $\tilde{Y}_0(x)=Y_{\bullet}(x)(\frac{1}{2}-B_{01}(x-))$, and we find $\tilde{Y}_0(\tau_1)=2(\frac{1}{2}-0)=1, \tilde{Y}_0(\tau_2)= 1(\frac{1}{2}-0)=\frac{1}{2}$, and $\tilde{Y}_0(\tau_3)=1(\frac{1}{2}-1)=-\frac{1}{2}$. The way we have defined $\tilde{\gamma}_{ij}$, seniorities with negative $\tilde{Y}_i$ do not contribute to the estimator, so we get a non-decreasing estimator for the integrated intensity, as we should. But the possible negativity of the "estimated number at risk" does suggest that the estimators (6.2) may behave badly in small samples. For large samples, they behave reasonably, however, as is seen from the following result.

<u>Theorem 6.1</u>. Assume that $P_i(\cdot)$ is bounded away from zero. Then

$$\sup_{x\in[0,z]} |\tilde{\gamma}_{ij}^{(n)}(x) - \gamma_{ij}(x)| \xrightarrow{P} 0 \quad \text{as} \quad n\to\infty.$$

<u>Proof</u>: According to Theorem 5.2 and (2.3),

$$\sup_{x\in[0,z]} |\tilde{Y}_i^{(n)}(x)/n - P_i(x)| \xrightarrow{P} 0 \quad \text{as} \quad n\to\infty , \tag{6.3}$$

and the remaining part of the proof follows as in Theorem 5.2 by applying Lenglart's (1977) inequality. □

It is possible to derive the asymptotic distributional properties of the $\{\tilde{\gamma}_{ij}\}$ by an argument similar to the one in Theorem 5.4, and the covariance structure may be calculated by using some of the results in Theorem 7.1 and Appendix C. We will not do this here, however, since for the applications we have in mind, the estimators proposed in our next section usually will be preferred.

7. Estimation of piecewise constant intensities. Asymptotic results

We assume in this section that the intensities are piecewise constant as they were in Section 3. If the cohort had been observed completely, we would then have estimated the $\{\alpha_{ijr}\}$ by the occurrence/exposure rates (cf. Hoem, 1976)

$$\hat{\alpha}_{ijr} = F_{ij}^{(n)}(a_r)/L_i^{(n)}(a_r), \tag{7.1}$$

where $F_{ij}^{(n)}(a_r) = K_{ij}^{(n)}(a_{r+1})-K_{ij}^{(n)}(a_r)$ and $L_i^{(n)}(a_r) = \int_{a_r}^{a_{r+1}} Y_i^{(n)}(u)du$.

For the situation considered in this paper, such detailed data are not available. However, we may "estimate" the exposure $L_i^{(n)}(a_r)$ by

$$\tilde{L}_i^{(n)}(a_r) = \int_{a_r}^{a_{r+1}} \tilde{Y}_i^{(n)}(u)du , \tag{7.2}$$

where $\tilde{Y}_i^{(n)}$ is given by (6.1). We are therefore lead to the estimators

$$\tilde{\alpha}_{ijr}^{(n)} = F_{ij}^{(n)}(a_r)/\tilde{L}_i^{(n)}(a_r) , \tag{7.3}$$

which are close to the estimators (3.3). By (6.3) it is seen that the estimators (7.3) are consistent, and we can also prove the following result.

<u>Theorem 7.1</u> Consider a fixed pair (i,j) and assume that

$\int_{a_r}^{a_{r+1}} P_i(u)du > 0$ for $r = 1,\dots,R$. Then,

$$\{\sqrt{n}(\tilde{\alpha}_{ijr}^{(n)}-\alpha_{ijr}), r=1,\dots,R\} \xrightarrow{\mathcal{D}} N_R(0,\Sigma),$$

where $\Sigma = (\sigma_{qr})$ is given by

$$\sigma_{rr} = \frac{\alpha_{ijr}}{\int_{a_r}^{a_{r+1}} P_i(\sigma)d\sigma}$$

$$+ \frac{2\alpha_{ijr}^2}{(\int_{a_r}^{a_{r+1}} P_i(\sigma)d\sigma)^2} \int_{a_r}^{a_{r+1}} \int_{a_r}^{r} \int_0^u \frac{p(u)p(r)}{p(\sigma)} \bar{P}_i(\sigma)(1-\bar{P}_i(\sigma))\mu(\sigma)d\sigma du dr \tag{7.4a}$$

and

$$\sigma_{qr} = \frac{\alpha_{ijq}\,\alpha_{ijr}}{\int_{a_q}^{a_{q+1}} P_i(\sigma)d\sigma \int_{a_r}^{a_{r+1}} P_i(\sigma)d\sigma} \int_{a_r}^{a_{r+1}} \int_{a_q}^{a_{q+1}} \int_0^u \frac{p(u)p(r)}{p(\sigma)} \bar{P}_i(\sigma)(1-\bar{P}_i(\sigma))\mu(\sigma)d\sigma du dr \tag{7.4b}$$

for $q < r$, and $N_R(0,\Sigma)$ denotes the R-dimensional multivariate normal distribution.

<u>Proof</u>. Let $X_{ij}^{(n)}$ be given as in Section 5 and introduce

$$V_i^{(n)}(t) = \sqrt{n}(\int_0^t \tilde{Y}_i^{(n)}(u)/n\,du - \int_0^t P_i(u)du).$$

Then, by a Taylor series expansion it follows that $\{\sqrt{n}(\tilde{\alpha}_{ijr}^{(n)}-\alpha_{ijr}),\ r=1,\dots,R\}$ has the same asymptotic distribution as the vector with r-th component

$$[\int_{a_r}^{a_{r+1}} P_i(u)du]^{-1}(X_{ij}^{(n)}(a_{r+1})-X_{ij}^{(n)}(a_r))$$
$$- \alpha_{ijr}[\int_{a_r}^{a_{r+1}} P_i(u)du]^{-1}(V_i^{(n)}(a_{r+1})-V_i^{(n)}(a_r)). \tag{7.5}$$

Now we may write

$$V_i^{(n)}(t) = \int_0^t \bar{P}_i(s)\sqrt{n}(Y_\bullet^{(n)}(s)/n-p(s))ds$$
$$+ \int_0^t p(s)\sqrt{n}(B_{\bullet i}^{(n)}(s-)-\beta_{\bullet i}(s))ds$$
$$- \int_0^t p(s)\sqrt{n}(B_{i\bullet}^{(n)}(s-)-\beta_{i\bullet}(s))ds$$
$$+ \sqrt{n}\left(\frac{N_i^{(n)}}{n} - \pi_i\right)\int_0^t p(s)ds$$
$$+ \int_0^t\left(B_{\bullet i}^{(n)}(s-)-B_{i\bullet}^{(n)}(s-)+ \frac{N_i^{(n)}}{n} - \bar{P}_i(s)\right)\sqrt{n}(Y_\bullet^{(n)}(s)/n-p(s))ds,$$

and it follows by a Skorohod construction that the sequence of processes $(X_{ij}^{(n)},V_i^{(n)})$ converges weakly to a limiting Gaussian process (X_{ij},V_i), where X_{ij} is given in Theorem 5.3 and

$$V_i(t) = \int_0^t \bar{P}_i(s)U(s)ds + \int_0^t p(s)\{Z_{\bullet i}(s)-Z_{i\bullet}(s)+M_i\}ds. \qquad (7.6)$$

Here U and Z_{ij} are given by (5.7) and (5.10), respectively, and M_i is a normal $N(0,\pi_i(1-\pi_i))$ random variable, independent of X_{ij},U,Z_{ij}, describing the number of individuals which start out in state $i \in \underline{\underline{K}}$ at seniority 0. Simple calculations show that

$$\mathrm{Cov}(X_{kl}(t),M_i) = \pi_i \int_0^t P_{ik}(0,s)\alpha_{kl}(s)ds - \pi_i\nu_{kl}(t) \qquad (7.7)$$

and

$$\mathrm{Cov}(U(t),M_i) = 0. \qquad (7.8)$$

Substituting X_{ij} and V_i for $X_{ij}^{(n)}$ and $V_i^{(n)}$ in (7.5) we get random variables having the same distribution as the asymptotic distribution of $\{\sqrt{n}(\tilde{\alpha}_{ijr}^{(n)}-\alpha_{ijr}),r=1,\ldots,R\}$.

The expressions for the variances and covariances follow by some straightforward, but very tedious calculations given in some detail in Appendix C. □

The first term in σ_{rr} is exactly the asymptotic variance of the occurrence/exposure rate (7.1). Therefore, the efficiency of our estimation method may easily be compared with the situation where complete information is available and the occurrence/exposure rates are used. An example of such efficiency calculations is given in the succeeding section. Note also that the asymptotic variances and covariances may be estimated consistently by substituting $\tilde{Y}_i(x)/n$ for $P_i(x)$, $B_{\cdot i}(x-)-B_{i\cdot}(x-)+N_i/n$ for $\bar{P}_i(x)$, $Y_{\cdot}(x)/n$ for $p(x)$, and $\tilde{\alpha}_{ijr}$ for α_{ijr} in (7.4a) and (7.4b). Moreover, by (7.4b) σ_{qr} is always positive, so that the estimators $(\tilde{\alpha}_{ijr}, r=1,\ldots,R)$ are all positively correlated.

8. An example. A first marriage model.

In order to illustrate the use of the estimators (7.3) based on the demographic incidence rates, and to compare them with the usual occurrence/exposure rates (7.1), we have studied a first marriage model. This simple Markov model is illustrated in Fig. 1. All women start out in state 0. Once a woman gets married she moves to state 1, where she remains until death. At death she moves on to state 2. A woman who dies before she gets married (for the first time), moves directly from state 0 to state 2.

For the female birth cohort of 32542 women born in Denmark in 1940, data were available so that we could estimate the first marriage intensty by the incidence rate method (7.3), as well as by the occurrence/exposure rates (7.1). (The data were also used by Finnäs, 1980, for illustrative purposes.) Let us assume that the intensities of the model in Fig. 1 are constant over single year age intervals. Denote the first marriage intensity and the force of mortality in the age interval $(r,r+1]$ by α_r and μ_r, respectively. Then the occurrence/exposure rates $\{\hat{\alpha}_r\}$, and the estimates based on the incidence rates $\{\tilde{\alpha}_r\}$, are given in Table 1 and shown in Fig. 2. The differences between the two sets of estimates are quite small.

We have also estimated the asymptotic variances and covariances given by (7.4). For the simple Markov model considered in this section, (7.4a) and (7.4b) are easily expressed as functions of the $\{\alpha_r\}$ and the $\{\mu_r\}$. This follows since the integrals in (7.4) may be computed by using the relations $\bar{P}_0(\sigma)=\bar{P}_0(r)\exp(-\alpha_r(\sigma-r))$, $p(\sigma)=p(r)\exp(-\mu_r(\sigma-r))$, and $P_0(\sigma)=\bar{P}_0(\sigma)p(\sigma)$ for $\sigma\in(r,r+1]$. Here $\bar{P}_0(r)=\exp(-\sum_{s=0}^{r-1}\alpha_s)$ and $p(r)=\exp(-\sum_{s=0}^{r-1}\mu_s)$. In Table 1 we have given the estimated standard deviations and the asymptotic efficiency of $\tilde{\alpha}_r$ relative to $\hat{\alpha}_r$. The values have been computed by substituting $\tilde{\alpha}_r$ for α_r and $\hat{\mu}_r=2(Y.^{(n)}(r)-Y.^{(n)}(r+1))/(Y.^{(n)}(r)+Y.^{(n)}(r+1))$ for μ_r in (7.4a). The estimated standard deviations of the two methods differ very little, implying an efficiency of the incidence rate method of 99.5 per cent or more for all ages.

The estimated correlation coefficients between $\tilde{\alpha}_q$ and $\tilde{\alpha}_r$ do not exceed 0.005 for any two age intervals, so the $\{\tilde{\alpha}_r\}$ are nearly asymptotically independent. This clarifies that the slightly lower estimates obtained by the incidence rate method, are not due to the positive correlation between the estimates, but, as suggested by Finnäs (1980), due to the fact that out-migration for this birth cohort mostly takes place among the unmarried women. Thus, the assumption about non-differential "mortality" (or more correct, the total effect of mortality and migration) is not completely satisfied.

The very high efficiencies obtained in this example, are partly explained by the low mortality for ages below 40 years in the cohort of Danish women born in 1940. The average yearly "mortality" rate is 1.7 per mille. Since the two estimation methods coincide when there is no mortality, it is not surprising that we get such high efficiencies in our case. To further study the effect of the mortality, we have also caluclated the asymptotic relative efficiencies for situations where the $\{\alpha_r\}$ remain unchanged, but the mortality is increased by a factor 2, 4, 6, 8, or 10 for all ages. In all these cases the lowest efficiency was obtained for age 29, where it attained the values 0.990, 0.981, 0.972, 0.964, and 0.957, respectively. So the relative efficiencies will exceed 95 per cent, even when we increase the mortality with 1000 per cent. This suggests that in this particular example, the level of the mortality does not influence the efficiency of the incidence rate method (7.3) very much.

To further explore how the efficiencies depend on the values of the $\{\alpha_r\}$ and $\{\mu_r\}$, we have made some additional numerical computations. Since some cohorts are wathced over their entire life span, and not only over a limited period as in the previous example, we have calculated relative efficiencies for a period of 70 years. For simplicity we assume in all these examples that $\alpha_r=\alpha$ and $\mu_r=\mu$ for all r, for some α and μ. The resulting efficiencies for α=0.05, 0.10, 0.15 and μ=0.001, 0.01 are shown in Fig. 3.

The efficiencies decrease with age for all the six cases considered, and the higher the values of α and μ, the faster is the decrease. The value of α seems to be of great importance for the efficiency. The efficiencies are above 90 per cent for ages below 20 years. For higher ages the efficiencies may be low. Consider for example the case where α=0.15, μ=0.001. Here the efficiencies decrease sharply after age 30 years. The reason seems to be that at age 30 years, there will only be 1 per cent of the original cohort left in state 0. Therefore (6.1) will be a poor "estimator" for the true number at risk, which again makes $\tilde{\alpha}_r$ an unreliable estimator for α_r.

To summarize, the efficiencies for the incidence rate method (7.3) for the model in Fig. 1 seem to be rather high for younger ages, unless the parameters $\{\alpha_r\}$ take very large values. The efficiencies may become smaller in the higher age groups, where the number of individuals at risk is small. The level of mortality is also important, but it does not seem to influence the efficiencies as much as do the values of the $\{\alpha_r\}$.

Acknowledgements

We are grateful to Jan M. Hoem for helpful comments and Martin Jacobsen for clarifying discussions. Ørnulf Borgan was supported by The Norwegian Research Council for Science and the Humanities, the Association of Norwegian Insurance Companies and Johan and Mimi Wessmann's legacy. Part of his work was done during a visit in the Department of Statistics at Stanford University.

Appendix A. Proof of Theorem 5.3

In this appendix, we prove Theorem 5.3, which also may be useful in other applications than those of the present paper. Theorem 5.3 is an ordinary central limit theorem in $D^m[0,1]$ for the number of transitions between the various states in a time continuous Markov chain. (Without loss of generality we may take $z=1$.) By Hoem and Aalen (1978, formula (12)), the covariance structure of the number of transitions is exactly our (5.6). By Hahn (1978, Theorem 2), therefore, to show tightness, it is sufficient to demonstrate that there exist nondecreasing continuous functions G and F on $[0,1]$ and numbers $\alpha>\frac{1}{2}$, $\beta>1$ such that for all $0\leq s\leq t\leq u\leq 1$ the following two conditions hold true

$$E[K_{ij}(t) - K_{ij}(s) - \int_s^t P_i(\sigma)\alpha_{ij}(\sigma)d\sigma]^2 \leq (G(t) - G(s))^\alpha \tag{A.1}$$

and

$$E[(K_{ij}(u)-K_{ij}(t)-\int_t^u P_i(\sigma)\alpha_{ij}(\sigma)d\sigma)^2(K_{ij}(t)-K_{ij}(s)-\int_s^t P_i(\sigma)\alpha_{ij}(\sigma)d\sigma)^2] \leq (F(u) - F(s))^\beta \tag{A.2}$$

for all $i,j\in \underline{\underline{I}}$, $i\neq j$, where $K_{ij}(t)$ now denotes the number of jumps in $[0,t]$ from state i directly to state j for a single individual. We will show that the conditions are satisfied with $\alpha=1$ and $\beta=2$. Fix i,j and introduce $N_{ij}(s,t)=K_{ij}(t)-K_{ij}(s)$ as the relevant number of transitions in $(s,t]$.

We first prove (A.1). By Hoem (1968)

$$E[N_{ij}(s,t)]^2 = \int_s^t P_i(\sigma)\alpha_{ij}(\sigma)d\sigma+2\int_s^t\int_\sigma^t P_i(\sigma)\alpha_{ij}(\sigma)P_{ji}(\sigma,\tau)\alpha_{ij}(\tau)d\tau d\sigma,$$

and therefore

$$\begin{aligned}E[N_{ij}(s,t) - \int_s^t P_i(\sigma)\alpha_{ij}(\sigma)d\sigma]^2 &= \int_s^t P_i(\sigma)\alpha_{ij}(\sigma)d\sigma \\ &+ 2\int_s^t\int_\sigma^t P_i(\sigma)\alpha_{ij}(\sigma)P_{ji}(\sigma,\tau)\alpha_{ij}(\tau)d\tau d\sigma, \\ &- (\int_s^t P_i(\sigma)\alpha_{ij}(\sigma)d\sigma)^2 \leq C(t-s) + C^2(t-s)^2 \leq (C+C^2)(t-s),\end{aligned}$$

since $t-s$ is bounded by 1. This proves (A.1).

We then prove (A.2) in a similar way. The left side of (A.2) is the sum of the following six terms:

$$E[(N_{ij}(s,t))^2(N_{ij}(t,u))^2], \tag{A.3}$$

$$E[(N_{ij}(s,t))^2](\int_t^u P_i(\sigma)\alpha_{ij}(\sigma)d\sigma)^2$$
$$+ E[(N_{ij}(t,u))^2](\int_s^t P_i(\sigma)\alpha_{ij}(\sigma)d\sigma)^2, \tag{A.4}$$

$$- 2E[N_{ij}(s,t)(N_{ij}(t,u))^2](\int_s^t P_i(\sigma)\alpha_{ij}(\sigma)d\sigma)$$
$$- 2E[N_{ij}(t,u)(N_{ij}(s,t))^2](\int_t^u P_i(\sigma)\alpha_{ij}(\sigma)d\sigma), \tag{A.5}$$

$$4E[N_{ij}(s,t)N_{ij}(t,u)](\int_s^t P_i(\sigma)\alpha_{ij}(\sigma)d\sigma)(\int_t^u P_i(\sigma)\alpha_{ij}(\sigma)d\sigma), \tag{A.6}$$

$$(\int_s^t P_i(\sigma)\alpha_{ij}(\sigma)d\sigma)^2(\int_t^u P_i(\sigma)\alpha_{ij}(\sigma)d\sigma)^2, \tag{A.7}$$

and

$$- 2E[N_{ij}(s,t)](\int_s^t P_i(\sigma)\alpha_{ij}(\sigma)d\sigma)(\int_t^u P_i(\sigma)\alpha_{ij}(\sigma)d\sigma)^2$$
$$- 2E[N_{ij}(t,u)](\int_t^u P_i(\sigma)\alpha_{ij}(\sigma)d\sigma)(\int_s^t P_i(\sigma)\alpha_{ij}(\sigma)d\sigma)^2. \tag{A.8}$$

The mixed moments of order 3 and 4 introduced in the terms above are given in (A.9) through (A.14) below. By adding the terms (A.3) to (A.8) and bounding all $\alpha_{ij}(\cdot)$ by the constant C and all probabilities by 1, it follows that there exists a constant K such that the left side of (A.2) is bounded by $K(u-s)^2$, and condition (A.2) is proved.

The higher order moments needed above may be found by the technique used by Hoem (1968). For completeness we also restate some of his results. Remember that we assume $0 \leq s \leq t \leq u \leq 1$. Then,

$$E[N_{ij}(s,t)] = \int_s^t P_i(\sigma)\alpha_{ij}(\sigma)d\sigma, \tag{A.9}$$

$$E[(N_{ij}(s,t)^2] = \int_s^t P_i(\sigma)\alpha_{ij}(\sigma)d\sigma$$
$$+ 2\int_s^t\int_\sigma^t P_i(\sigma)\alpha_{ij}(\sigma)P_{ji}(\sigma,\tau)\alpha_{ij}(\tau)d\tau d\sigma, \tag{A.10}$$

$$E[N_{ij}(s,t)N_{ij}(t,u)] = \int_s^t\int_t^u P_i(\sigma)\alpha_{ij}(\sigma)P_{ji}(\sigma,\tau)\alpha_{ij}(\tau)d\tau d\sigma, \tag{A.11}$$

$$E[(N_{ij}(s,t))^2 N_{ij}(t,u)] = \int_s^t\int_t^u P_i(\sigma)\alpha_{ij}(\sigma)P_{ji}(\sigma,\tau)\alpha_{ij}(\tau)d\tau d\sigma$$
$$+ 2\int_s^t\int_\sigma^t\int_t^u P_i(\sigma)\alpha_{ij}(\sigma)P_{ji}(\sigma,\tau)\alpha_{ij}(\tau)P_{ji}(\tau,\nu)\alpha_{ij}(\nu)d\nu d\tau d\sigma, \tag{A.12}$$

$$E[N_{ij}(s,t)(N_{ij}(t,u))^2] = \int_s^t\int_t^u P_i(\sigma)\alpha_{ij}(\sigma)P_{ji}(\sigma,\tau)\alpha_{ij}(\tau)d\tau d\sigma$$
$$+ 2\int_s^t\int_t^u\int_\tau^u P_i(\sigma)\alpha_{ij}(\sigma)P_{ji}(\sigma,\tau)\alpha_{ij}(\tau)P_{ji}(\tau,\nu)\alpha_{ij}(\nu)d\nu d\tau d\sigma, \tag{A.13}$$

and

$$E[(N_{ij}(s,t))^2(N_{ij}(t,u))^2] = \int_s^t\int_t^u P_i(\sigma)\alpha_{ij}(\sigma)P_{ji}(\sigma,\tau)\alpha_{ij}(\tau)d\tau d\sigma$$
$$+ 2\int_s^t\int_\sigma^t\int_t^u P_i(\sigma)\alpha_{ij}(\sigma)P_{ji}(\sigma,\tau)\alpha_{ij}(\tau)P_{ji}(\tau,\nu)\alpha_{ij}(\nu)d\nu d\tau d\sigma$$
$$+ 2\int_s^t\int_t^u\int_\tau^u P_i(\sigma)\alpha_{ij}(\sigma)P_{ji}(\sigma,\tau)\alpha_{ij}(\tau)P_{ji}(\tau,\nu)\alpha_{ij}(\nu)d\nu d\tau d\sigma \tag{A.14}$$
$$+4\int_s^t\int_\sigma^t\int_t^u\int_\nu^u P_i(\sigma)\alpha_{ij}(\sigma)P_{ji}(\sigma,\tau)\alpha_{ij}(\tau)P_{ji}(\tau,\nu)\alpha_{ij}(\nu)P_{ji}(\nu,\mu)\alpha_{ij}(\mu)d\mu d\nu d\tau d\sigma.$$

Appendix B. The covariance structure of the limiting process in Theorem 5.4

By (5.10) the covariance between $Z_{ij}(x)$ and $Z_{kl}(y)$ for $x \leq y$ consists of the sum of the following four terms:

$$\mathrm{Cov}\left(-\int_0^x U(s)[p(s)]^{-2}d\nu_{ij}(s),\ -\int_0^y U(t)[p(t)]^{-2}d\nu_{kl}(t)\right), \tag{B.1}$$

$$\mathrm{Cov}\left(-\int_0^x U(s)[p(s)]^{-2}d\nu_{ij}(s),\ X_{kl}(y)[p(y)]^{-1} - \int_0^y X_{kl}(t)\mu(t)[p(t)]^{-1}dt\right), \tag{B.2}$$

$$\mathrm{Cov}\left(X_{ij}(x)[p(x)]^{-1} - \int_0^x X_{ij}(s)\mu(s)[p(s)]^{-1}ds,\ -\int_0^y U(t)[p(t)]^{-2}d\nu_{kl}(t)\right), \tag{B.3}$$

and

$$\mathrm{Cov}\left(X_{ij}(x)[p(x)]^{-1} - \int_0^x X_{ij}(s)\mu(s)[p(s)]^{-1}ds,\ X_{kl}(y)[p(y)]^{-1} - \int_0^y X_{kl}(t)\mu(t)[p(t)]^{-1}dt\right). \tag{B.4}$$

By (5.9) and integration by parts, (B.1) equals

$$\int_0^x\int_0^u [p(r)]^{-2}[p(u)]^{-1}d\nu_{ij}(r)d\nu_{kl}(u) + \int_0^x\int_0^u [p(r)]^{-2}[p(u)]^{-1}d\nu_{kl}(r)d\nu_{ij}(u)$$
$$+ \int_0^x [p(r)]^{-2}d\nu_{ij}(r)\int_x^y [p(u)]^{-1}d\nu_{kl}(u) - \int_0^x [p(u)]^{-1}d\nu_{ij}(u)\int_0^y [p(r)]^{-1}d\nu_{kl}(r). \tag{B.5}$$

Also by integration by parts, (B.2) may be rewritten as

$$-\int_0^x \Big[\mathrm{Cov}(U(s),X_{kl}(y))(p(y))^{-1}$$
$$- \int_0^y \mathrm{Cov}(U(s),X_{kl}(t))\mu(t)(p(t))^{-1}dt\Big](p(s))^{-2}d\nu_{ij}(s)$$
$$= -\int_{s=0}^{x}\int_{t=0}^{y} \mathrm{Cov}(U(s),X_{kl}(dt))(p(t))^{-1}(p(s))^{-2}d\nu_{ij}(s),$$

which by (5.8) equals

$$\int_0^x [p(r)]^{-1} d\nu_{ij}(r) \int_0^y [p(u)]^{-1} d\nu_{kl}(u) - \int_0^x\int_0^u [p(r)]^{-2}[p(u)]^{-1} d\nu_{ij}(r) d\nu_{kl}(u)$$
$$- \int_0^x\int_0^u [p(r)]^{-2}[p(u)]^{-1} d\nu_{kl}(r) d\nu_{ij}(u) - \int_0^x [p(r)]^{-2} d\nu_{ij}(r) \int_x^y [p(u)]^{-1} d\nu_{kl}(u). \qquad (B.6)$$

By quite parallel calculations, (B.3) is also found to equal (B.6). Finally, by integration by parts, (B.4) equals

$$\int_0^x [p(y)p(s)]^{-1} \operatorname{Cov}(X_{ij}(ds), X_{kl}(y))$$
$$- \int_{t=0}^{y} \int_{s=0}^{x} [p(t)p(s)]^{-1} \mu(t) \operatorname{Cov}(X_{ij}(ds), X_{kl}(t)) dt.$$

By (5.6), $\operatorname{Cov}(X_{ij}(ds), X_{kl}(t))$ equals

$$\int_s^t P_{jk}(s,\tau)\alpha_{kl}(\tau) d\tau d\nu_{ij}(s) + \int_0^s P_{li}(\tau,s)\alpha_{ij}(s) d\nu_{kl}(\tau) ds$$
$$+ \delta_{ik}\delta_{jl} d\nu_{ij}(s) - \nu_{kl}(t) d\nu_{ij}(s)$$

for $s \leq t$, and it equals

$$\int_0^t P_{li}(\sigma,s)\alpha_{ij}(s) d\nu_{kl}(\sigma) ds - \nu_{kl}(t) d\nu_{ij}(s)$$

for $s > t$. Therefore, by some straightforward calculations, (B.4) is found equal

$$\int_0^x\int_0^u P_{jk}(r,u)\alpha_{kl}(u)[p(r)p(u)]^{-1} d\nu_{ij}(r) du$$
$$+ \int_0^x\int_0^r P_{li}(u,r)\alpha_{ij}(r)[p(r)p(u)]^{-1} d\nu_{kl}(u) dr$$
$$+ \int_x^y\int_0^x P_{jk}(r,u)\alpha_{kl}(u)[p(r)p(u)]^{-1} d\nu_{ij}(r) du + \delta_{ik}\delta_{jl}\int_0^x [p(u)]^{-2} d\nu_{ij}(u)$$
$$- \int_0^x [p(r)]^{-1} d\nu_{ij}(r) \int_0^y [p(u)]^{-1} d\nu_{kl}(u). \qquad (B.7)$$

Adding together (B.5), two times (B.6), and (B.7), we get the covariance structure (5.11).

Appendix C. The variances and covariances in Theorem 7.1

The computation of the variances and covariances in Theorem 7.1 is straightforward, but lengthy and very tedious. So we do only give the main steps in the calculations and the most important intermediate results. By (7.5) it is seen that to compute the asymptotic variances and covariances we need expressions for the following terms:

$$\mathrm{Cov}(X_{ij}(a_{q+1})-X_{ij}(a_q),X_{ij}(a_{r+1})-X_{ij}(a_r)), \tag{C.1}$$

$$\mathrm{Cov}(V_i(a_{q+1})-V_i(a_q),V_i(a_{r+1})-V_i(a_r)), \tag{C.2}$$

and

$$\mathrm{Cov}(X_{ij}(a_{q+1})-X_{ij}(a_q),V_i(a_{r+1})-V_i(a_r)). \tag{C.3}$$

By (5.6), (C.1) equals

$$\alpha_{ijr}\int_{a_r}^{a_{r+1}} P_i(\sigma)d\sigma + 2\alpha_{ijr}^2\int_{a_r}^{a_{r+1}}\int_{a_r}^{\tau} P_i(\sigma)P_{ji}(\sigma,\tau)d\sigma d\tau - \alpha_{ijr}^2\left(\int_{a_r}^{a_{r+1}} P_i(\sigma)d\sigma\right)^2 \tag{C.4}$$

for $q = r$, and it equals

$$\alpha_{ijq}\,\alpha_{ijr}\int_{a_r}^{a_{r+1}}\int_{a_q}^{a_{q+1}} P_i(\sigma)P_{ji}(\sigma,\tau)d\sigma d\tau - \alpha_{ijq}\,\alpha_{ijr}\int_{a_q}^{a_{q+1}} P_i(\sigma)d\sigma \int_{a_r}^{a_{r+1}} P_i(\sigma)d\sigma \tag{C.5}$$

for $q < r$.

To compute (C.2), note first that by (7.6) it equals the sum of the following nine terms:

$$\mathrm{Cov}\left(\int_{a_q}^{a_{q+1}} \bar{P}_i(\sigma)U(\sigma)d\sigma,\ \mathrm{Cov}\left(\int_{a_r}^{a_{r+1}} \bar{P}_i(\sigma)U(\sigma)d\sigma\right),\right. \tag{C.6}$$

$$\mathrm{Cov}\left(\int_{a_q}^{a_{q+1}} \bar{P}_i(\sigma)U(\sigma)d\sigma,\ \int_{a_r}^{a_{r+1}} p(\sigma)[Z_{\cdot i}(\sigma)-Z_{i\cdot}(\sigma)]d\sigma\right), \tag{C.7}$$

$$\mathrm{Cov}\left(\int_{a_q}^{a_{q+1}} \bar{P}_i(\sigma)U(\sigma)d\sigma,\ M_i\int_{a_r}^{a_{r+1}} p(\sigma)d\sigma\right), \tag{C.8}$$

$$\operatorname{Cov}\left(\int_{a_q}^{a_{q+1}} p(\sigma)[Z_{\cdot i}(\sigma)-Z_{i\cdot}(\sigma)]d\sigma,\ \int_{a_r}^{a_{r+1}} \bar{P}_i(\sigma)U(\sigma)d\sigma\right), \tag{C.9}$$

$$\operatorname{Cov}\left(\int_{a_q}^{a_{q+1}} p(u)[Z_{\cdot i}(u)-Z_{i\cdot}(u)]du,\ \int_{a_r}^{a_{r+1}} p(r)[Z_{\cdot i}(r)-Z_{i\cdot}(r)]dr\right), \tag{C.10}$$

$$\operatorname{Cov}\left(\int_{a_q}^{a_{q+1}} p(\sigma)[Z_{\cdot i}(\sigma)-Z_{i\cdot}(\sigma)]d\sigma,\ M_i\int_{a_r}^{a_{r+1}} p(\sigma)d\sigma\right), \tag{C.11}$$

$$\operatorname{Cov}\left(M_i\int_{a_q}^{a_{q+1}} p(\sigma)d\sigma,\ \int_{a_r}^{a_{r+1}} \bar{P}_i(\sigma)U(\sigma)d\sigma\right), \tag{C.12}$$

$$\operatorname{Cov}\left(M_i\int_{a_q}^{a_{q+1}} p(\sigma)d\sigma,\ \int_{a_r}^{a_{r+1}} p(\sigma)[Z_{\cdot i}(\sigma)-Z_{i\cdot}(\sigma)]d\sigma\right), \tag{C.13}$$

and

$$\operatorname{Cov}\left(M_i\int_{a_q}^{a_{q+1}} p(\sigma)d\sigma,\ M_i\int_{a_r}^{a_{r+1}} p(\sigma)d\sigma\right). \tag{C.14}$$

By (2.1) and (5.9), (C.6) equals

$$2\int_{a_r}^{a_{r+1}}\int_{a_r}^{\tau} \bar{P}_i(\sigma)P_i(\tau)d\sigma d\tau - \left(\int_{a_r}^{a_{r+1}} P_i(\sigma)d\sigma\right)^2 \tag{C.15}$$

for $q = r$, and it equals

$$\left[\int_{a_q}^{a_{q+1}} \bar{P}_i(\sigma)d\sigma - \int_{a_q}^{a_{q+1}} P_i(\sigma)d\sigma\right]\int_{a_r}^{a_{r+1}} P_i(\tau)d\tau \tag{C.16}$$

for $q < r$.

By (5.10) and calculations similar to those of Appendix B, we find that $\operatorname{Cov}(U(\sigma),Z_{kl}(\tau)) = 0$ for all σ,τ and all $k,l \in \underline{\underline{K}}$, $k \neq l$. Thus (C.7) and (C.9) both equal zero. Applying (2.1), (2.3), the fact that

$$d\nu_{ij}(\sigma) = p(\sigma)d\beta_{ij}(\sigma), \tag{C.17}$$

and the Kolmogorov differential equation

$$\frac{\partial}{\partial\tau}\bar{P}_{ji}(\sigma,\tau) = -\bar{P}_{ji}(\sigma,\tau)\alpha_{i\cdot}(\tau) + \sum_{k\in\underline{\underline{K}}\setminus\{i\}} \bar{P}_{jk}(\sigma,\tau)\alpha_{ki}(\tau), \tag{C.18}$$

it follows by (5.11) and some straightforward calculations using integration by parts, that

$$\begin{aligned}
&\operatorname{Cov}(Z_{\bullet i}(u)-Z_{i\bullet}(u),\ Z_{\bullet i}(r)-Z_{i\bullet}(r)) \\
&= \int_0^u [\bar{P}_{ii}(\sigma,u)+\bar{P}_{ii}(\sigma,r)]\frac{d\beta_{\bullet i}(\sigma)}{p(\sigma)} \\
&- \int_0^u [1+\bar{P}_i(u)+\bar{P}_i(r)-2\bar{P}_i(\sigma)]\frac{d\bar{P}_i(\sigma)}{p(\sigma)} \\
&- \sum_{k\in\underline{\underline{K}}\setminus\{i\}} \int_0^u [\bar{P}_{ki}(\sigma,u)+\bar{P}_{ki}(\sigma,r)]\frac{d\beta_{ik}(\sigma)}{p(\sigma)}
\end{aligned} \tag{C.19}$$

for $u \leqslant r$. Using this and (2.1), (C.10) is found to equal

$$\begin{aligned}
&2\int_{a_r}^{a_{r+1}}\int_{a_r}^{r}\int_0^u p(r)P_{ii}(\sigma,u)d\beta_{\bullet i}(\sigma)dudr \\
&+ 2\int_{a_r}^{a_{r+1}}\int_{a_r}^{r}\int_0^u p(u)P_{ii}(\sigma,r)d\beta_{\bullet i}(\sigma)dudr \\
&- 2\sum_{k\in\underline{\underline{K}}\setminus\{i\}}\int_{a_r}^{a_{r+1}}\int_{a_r}^{r}\int_0^u p(r)P_{ki}(\sigma,u)d\beta_{ik}(\sigma)dudr \\
&- 2\sum_{k\in\underline{\underline{K}}\setminus\{i\}}\int_{a_r}^{a_{r+1}}\int_{a_r}^{r}\int_0^u p(u)P_{ki}(\sigma,r)d\beta_{ik}(\sigma)dudr \\
&- 2\int_{a_r}^{a_{r+1}}\int_{a_r}^{r}\int_0^u \frac{p(u)p(r)}{p(\sigma)}[1+\bar{P}_i(u)+\bar{P}_i(r)-2\bar{P}_i(\sigma)]d\bar{P}_i(\sigma)dudr
\end{aligned} \tag{C.20}$$

for $q = r$, and it equals

$$\begin{aligned}
&\int_{a_r}^{a_{r+1}}\int_{a_q}^{a_{q+1}}\int_0^u p(r)P_{ii}(\sigma,u)d\beta_{\bullet i}(\sigma)dudr \\
&+ \int_{a_r}^{a_{r+1}}\int_{a_q}^{a_{q+1}}\int_0^u p(u)P_{ii}(\sigma,r)d\beta_{\bullet i}(\sigma)dudr \\
&- \sum_{k\in\underline{\underline{K}}\setminus\{i\}}\int_{a_r}^{a_{r+1}}\int_{a_q}^{a_{q+1}}\int_0^u p(r)P_{ki}(\sigma,u)d\beta_{ik}(\sigma)dudr \\
&- \sum_{k\in\underline{\underline{K}}\setminus\{i\}}\int_{a_r}^{a_{r+1}}\int_{a_q}^{a_{q+1}}\int_0^u p(u)P_{ki}(\sigma,r)d\beta_{ik}(\sigma)dudr
\end{aligned} \tag{C.21}$$

$$- \int_{a_r}^{a_{r+1}} \int_{a_q}^{a_{q+1}} \int_0^u \frac{p(u)p(r)}{p(\sigma)}[1+\bar{P}_i(u)+\bar{P}_i(r)-2\bar{P}_i(\sigma)]d\bar{P}_i(\sigma)dudr$$

for $q < r$.

Using (7.8), we see that (C.8) and (C.12) both equal zero. Furthermore, by (5.10), (7.7), (7.8) and integration by parts

$$\mathrm{Cov}(Z_{kl}(t),M_i) = \pi_i \int_0^t \bar{P}_{ik}(0,s)\alpha_{kl}(s)ds-\pi_i\beta_{kl}(t),$$

and it follows by (2.3) that

$$\mathrm{Cov}(Z_{\cdot i}(u)-Z_{i\cdot}(u),M_i) = \pi_i(\bar{P}_{ii}(0,u)-\bar{P}_i(u))-\pi_i(1-\pi_i).$$

Using this (C.11) is found to equal

$$\pi_i \int_{a_r}^{a_{r+1}} p(\sigma)d\sigma \int_{a_q}^{a_{q+1}} (P_{ii}(0,u)-P_i(u))du-\pi_i(1-\pi_i) \int_{a_q}^{a_{q+1}} p(\sigma)d\sigma \int_{a_r}^{a_{r+1}} p(\sigma)d\sigma, \quad (C.22)$$

and (C.13) is given by a similar expression.

Finally, of course, (C.14) equals

$$\pi_i(1-\pi_i) \int_{a_q}^{a_{q+1}} p(\sigma)d\sigma \int_{a_r}^{a_{r+1}} p(\sigma)d\sigma. \quad (C.23)$$

Now, by (7.6), (C.3) is the sum of the following three terms:

$$\mathrm{Cov}(X_{ij}(a_{q+1})-X_{ij}(a_q), \int_{a_r}^{a_{r+1}} \bar{P}_i(\sigma)U(\sigma)d\sigma), \quad (C,24)$$

$$\mathrm{Cov}(X_{ij}(a_{q+1})-X_{ij}(a_q), \int_{a_r}^{a_{r+1}} p(r)[Z_{\cdot i}(r)-Z_{i\cdot}(r)]dr) \quad (C.25)$$

and

$$\mathrm{Cov}(X_{ij}(a_{q+1})-X_{ij}(a_q), M_i \int_{a_r}^{a_{r+1}} p(\sigma)d\sigma). \quad (C.26)$$

Using (5.8), we find that (C.24) equals

$$2\alpha_{ijr} \int_{a_r}^{a_{r+1}} \int_{a_r}^{\tau} \bar{P}_i(\sigma)P_i(\tau)d\sigma d\tau - \alpha_{ijr}(\int_{a_r}^{a_{r+1}} P_i(\sigma)d\sigma)^2 \quad (C.27)$$

for $r = q$; it equals

$$\alpha_{ijq} \int_{a_r}^{a_{r+1}} P_i(\sigma)d\sigma \int_{a_q}^{a_{q+1}} \bar{P}_i(\tau)d\tau - \alpha_{ijq} \int_{a_r}^{a_{r+1}} P_i(\sigma)d\sigma \int_{a_q}^{a_{q+1}} P_i(\tau)d\tau \qquad (C.28)$$

for $q < r$; and it equals

$$\alpha_{ijq} \int_{a_r}^{a_{r+1}} \bar{P}_i(\sigma)d\sigma \int_{a_q}^{a_{q+1}} P_i(\tau)d\tau - \alpha_{ijq} \int_{a_r}^{a_{r+1}} P_i(\sigma)d\sigma \int_{a_q}^{a_{q+1}} P_i(\tau)d\tau \qquad (C.29)$$

for $q > r$.

By (5.10) and calculations similar to those of Appendix B we find, using (2.3), (C.17) and (C.18), that

$$\mathrm{Cov}(X_{ij}(s),\ Z_{\cdot i}(r)-Z_{i\cdot}(r))$$

equals

$$\begin{aligned}
&\int_0^r \int_\sigma^s P_{ii}(\sigma,u)\alpha_{ij}(u)dud\beta_{i\cdot}(\sigma) \\
&- \sum_{k\in\underline{\underline{K}}\setminus\{i\}} \int_0^r \int_\sigma^s P_{ki}(\sigma,u)\alpha_{ij}(u)dud\beta_{ik}(\sigma) \\
&+ \int_0^r \bar{P}_{ji}(\sigma,r)d\beta_{ij}(\sigma) + \int_0^r \frac{\nu_{ij}(u)}{p(u)}\, d\bar{P}_i(u) \qquad (C.30) \\
&- \beta_{ij}(r) - \int_0^r \beta_{ij}(u)d\bar{P}_i(u) \\
&- \nu_{ij}(s) \int_0^r \frac{1}{p(u)}\, d\bar{P}_i(u)
\end{aligned}$$

for $s \geq r$, and it equals

$$\begin{aligned}
&\int_0^s \int_\sigma^s P_{ii}(\sigma,u)\alpha_{ij}(u)dud\beta_{\cdot i}(\sigma) \\
&- \sum_{k\in\underline{\underline{K}}\setminus\{i\}} \int_0^s \int_\sigma^s P_{ki}(\sigma,u)\alpha_{ij}(u)dud\beta_{ik}(\sigma) \\
&+ \int_0^s \bar{P}_{ji}(\sigma,r)d\beta_{ij}(\sigma) + \int_0^s \frac{\nu_{ij}(u)}{p(u)}\, d\bar{P}_i(u) \qquad (C.31) \\
&- \beta_{ij}(s) - \int_0^s \beta_{ij}(u)d\bar{P}_i(u) - \nu_{ij}(s) \int_0^s \frac{d\bar{P}_i(u)}{p(u)} \\
&- \beta_{ij}(s)\bar{P}_i(r) + \beta_{ij}(s)\bar{P}_i(s)
\end{aligned}$$

for $s < r$. Using this (C.25) is found to equal

$$\alpha_{ijr}\int_{a_r}^{a_{r+1}}\int_{a_r}^{r}[P_{ji}(\sigma,r)P_i(\sigma)-P_i(r)\bar{P}_i(\sigma)-\bar{P}_i(\sigma)p(r)+\bar{P}_i^2(\sigma)p(r)]d\sigma dr$$
$$-\alpha_{ijr}\int_{a_r}^{a_{r+1}}\int_{a_r}^{r}\int_{u}^{a_{r+1}}\frac{P_i(\sigma)p(r)}{p(u)}d\sigma d\bar{P}_i(u)dr$$
$$-\alpha_{ijr}\int_{a_r}^{a_{r+1}}P_i(u)du\int_{a_r}^{a_{r+1}}p(r)dr\int_0^{a_r}\frac{d\bar{P}_i(u)}{p(u)} \tag{C.32}$$
$$+\int_{a_r}^{a_{r+1}}\int_{a_r}^{r}\int_0^{u}p(r)P_{ii}(\sigma,u)d\beta_{\bullet i}(\sigma)dudr$$
$$+\alpha_{ijr}\int_{a_r}^{a_{r+1}}\int_{r}^{a_{r+1}}\int_0^{r}p(r)P_{ii}(\sigma,u)d\beta_{\bullet i}(\sigma)dudr$$
$$-\alpha_{ijr}\sum_{k\in\underline{\underline{K}}\setminus\{i\}}\int_{a_r}^{a_{r+1}}\int_{a_r}^{r}\int_0^{u}p(r)P_{ki}(\sigma,u)d\beta_{ik}(\sigma)dudr$$
$$-\alpha_{ijr}\sum_{k\in\underline{\underline{K}}\setminus\{i\}}\int_{a_r}^{a_{r+1}}\int_{r}^{a_{r+1}}\int_0^{r}p(r)P_{ki}(\sigma,u)d\beta_{ik}(\sigma)dudr$$

for $q = r$; it equals

$$\alpha_{ijq}\int_{a_r}^{a_{r+1}}\int_{a_q}^{a_{q+1}}[P_{ji}(\sigma,r)P_i(\sigma)-P_i(r)\bar{P}_i(\sigma)-\bar{P}_i(\sigma)p(r)+\bar{P}_i{}^2(\sigma)p(r)]d\sigma dr$$
$$-\alpha_{ijq}\int_{a_r}^{a_{r+1}}\int_{a_q}^{a_{q+1}}\int_{u}^{a_{q+1}}\frac{P_i(\sigma)p(r)}{p(u)}d\sigma d\bar{P}_i(u)dr \tag{C.33}$$
$$-\alpha_{ijq}\int_{a_q}^{a_{q+1}}P_i(u)du\int_{a_r}^{a_{r+1}}p(r)dr\int_0^{a_q}\frac{d\bar{P}_i(u)}{p(u)}$$
$$+\alpha_{ijq}\int_{a_r}^{a_{r+1}}\int_{a_q}^{a_{q+1}}\int_0^{u}p(r)P_{ii}(\sigma,u)d\beta_{\bullet i}(\sigma)dudr$$
$$-\alpha_{ijq}\sum_{k\in\underline{\underline{K}}\setminus\{i\}}\int_{a_r}^{a_{r+1}}\int_{a_q}^{a_{q+1}}\int_0^{u}p(r)P_{ki}(\sigma,u)d\beta_{ik}(\sigma)dudr$$

for $q < r$, and it equals

$$\alpha_{ijq} \int_{a_r}^{a_{r+1}} \int_{a_q}^{a_{q+1}} \int_0^r p(r)P_{ii}(\sigma,u)d\beta_{\cdot i}(\sigma)dudr$$

$$- \alpha_{ijq} \sum_{k\in\underline{\underline{K}}\setminus\{i\}} \int_{a_r}^{a_{r+1}} \int_{a_q}^{a_{q+1}} p(r)P_{ki}(\sigma,u)d\beta_{ik}(\sigma)dudr \tag{C.34}$$

$$- \alpha_{ijq} \int_{a_q}^{a_{q+1}} P_i(\sigma)d\sigma \int_{a_r}^{a_{r+1}} \int_0^r \frac{p(r)}{p(u)} d\bar{P}_i(u)dr$$

for $q > r$.

Finally, using (7.7), (C.26) is found to equal

$$\alpha_{ijq} \pi_i \int_{a_r}^{a_{r+1}} p(\sigma)d\sigma \int_{a_q}^{a_{q+1}} (P_{ii}(0,s)-P_i(s))ds \ . \tag{C.35}$$

Combining all this, and using the fact that $\int_0^u (1-2\bar{P}_i(\sigma))d\bar{P}_i(\sigma) = \bar{P}_i(u)(1-\bar{P}_i(u)) - \pi_i(1-\pi_i)$, the variances and covariances in Theorem 7.1 result.

REFERENCES

Aalen, O.O. (1978). Nonparametric inference for a family of counting processes. Ann.Statist. 6, 701-726.

Aalen, O.O. and Johansen, S. (1978). An empirical transition matrix for non-homogeneous Markov chains based on censored observations. Scand.J.Statist. 5, 141-150.

Andersen, P. K. and Gill, R. D. (1982). Cox's regression model for counting processes: a large sample study. Ann. Statist. 10, 1100-1120.

Billingsley, D. (1968). Convergence of probability measures. Wiley, New York.

Breslow, N. and Crowley, J. (1974). A large sample study of the life table and product-limit estimates under random censorship. Ann.Statist. 2, 437-453.

Finnäs, F. (1980). A method to estimate demographic intensities via cumulative incidence rates. Theor.Popul.Biol. 17, 365-379.

Gill, R.D. (1980). Censoring and stochastic integrals. Mathematical Centre Tracts 124, Mathematisch Centrum, Amsterdam.

Gill, R.D. (1981). Large sample behaviour of the product-limit estimator on the whole line. Preprint SW 74/81, Mathematisch Centrum, Amsterdam.

Hahn, M. (1978). Central limit theorems in D[0,1]. Z. Wahrscheinlichkeitstheorie verw. Gebiete 44, 89-101.

Hoem. J.M. (1968). Application of time-continuous Markov chains to life insurance. Memorandum 29/4-68, Inst. of Econ., Univ. of Oslo.

Hoem, J.M. (1969). Purged and partial Markov chains. Scand.Aktuar. Tidskr. 52, 147-155.

Hoem, J.M. (1976). The statistical theory of demographic rates. A review of current developments (with discussion). Scand.J.Statist. 3, 169-185.

Hoem, J.M. (1978). Demographic incidence rates. Theor.Popul.Biol. 14, 329-337.

Hoem, J.M. and Aalen, O.O. (1978). Actuarial values of payment streams. Scand.Actuarial J., 38-47.

Hoem, J.M. and Jensen, U.F. (1982). Multistate life table methodology: A probabilist critique. To appear in Multidimensional Mathematical Demography (Academic Press), edited by Kenneth C. Land and Andrei Rogers.

Jacobsen, M. (1982). Statistical analysis of counting processes. Lecture Notes in Statistics 12, Springer-Verlag, New York.

Lenglart, E. (1977). Relation de domination entre deux processus. Ann.Inst.Henri Poincaré 13, 171-179.

Rebolledo, R. (1978). Sur les applications de la théorie des martingales a l'étude statistique d'une famille de processus ponctueles. Springer Lecture Notes in Math. 636, 27-70.

Rebolledo, R. (1980). Central limit theorems for local martingales. Z. Wahrscheinlichkeitstheorie Verw. Gebiete 51, 269-286.

Table 1. Estimated first marriage intensities with standard deviations for women born in Denmark in 1940.

Age	(1) Occ/exp rates $\hat{\alpha}_r$ per 1000	(2) Incidence rate method $\tilde{\alpha}_r$ per 1000	(3) Standard deviation on $\hat{\alpha}_r$ per 1000	(4) Standard deviation on $\tilde{\alpha}_r$ per 1000	(5) Efficiency $\{(3)/(4)\}^2$
15	0.4	0.4	0.111	0.111	1.000
16	6.7	6.7	0.455	0.455	1.000
17	25.6	25.6	0.900	0.900	1.000
18	92.6	92.3	1.766	1.766	1.000
19	138.6	138.4	2.294	2.294	1.000
20	191.4	190.6	2.925	2.925	1.000
21	255.2	254.1	3.777	3.778	0.999
22	297.4	292.4	4.648	4.651	0.999
23	301.9	293.9	5.399	5.405	0.998
24	271.4	266.9	5.929	5.939	0.997
25	250.2	241.8	6.417	6.430	0.996
26	194.5	190.9	6.357	6.371	0.996
27	160.2	157.6	6.306	6.320	0.996
28	128.0	125.6	6.051	6.063	0.996
29	130.6	126.8	6.482	6.499	0.995
30	85.7	81.9	5.494	5.504	0.996
31	60.2	56.7	4.733	4.739	0.997
32	42.7	39.3	4.034	4.038	0.998
33	54.1	49.4	4.622	4.628	0.997
34	47.5	43.0	4.415	4.420	0.998
35	44.2	40.2	4.362	4.367	0.998
36	26.6	24.2	3.441	3.443	0.999
37	36.5	33.1	4.082	4.087	0.998
38	19.7	17.8	3.033	3.034	0.999

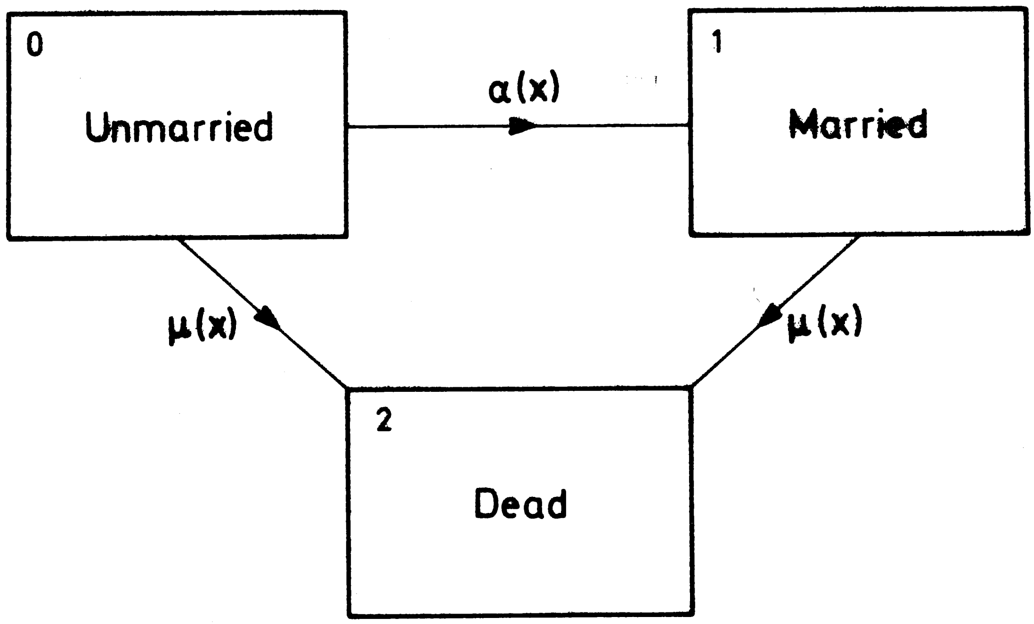

Fig. 1. A first marriage model for a female birth cohort.

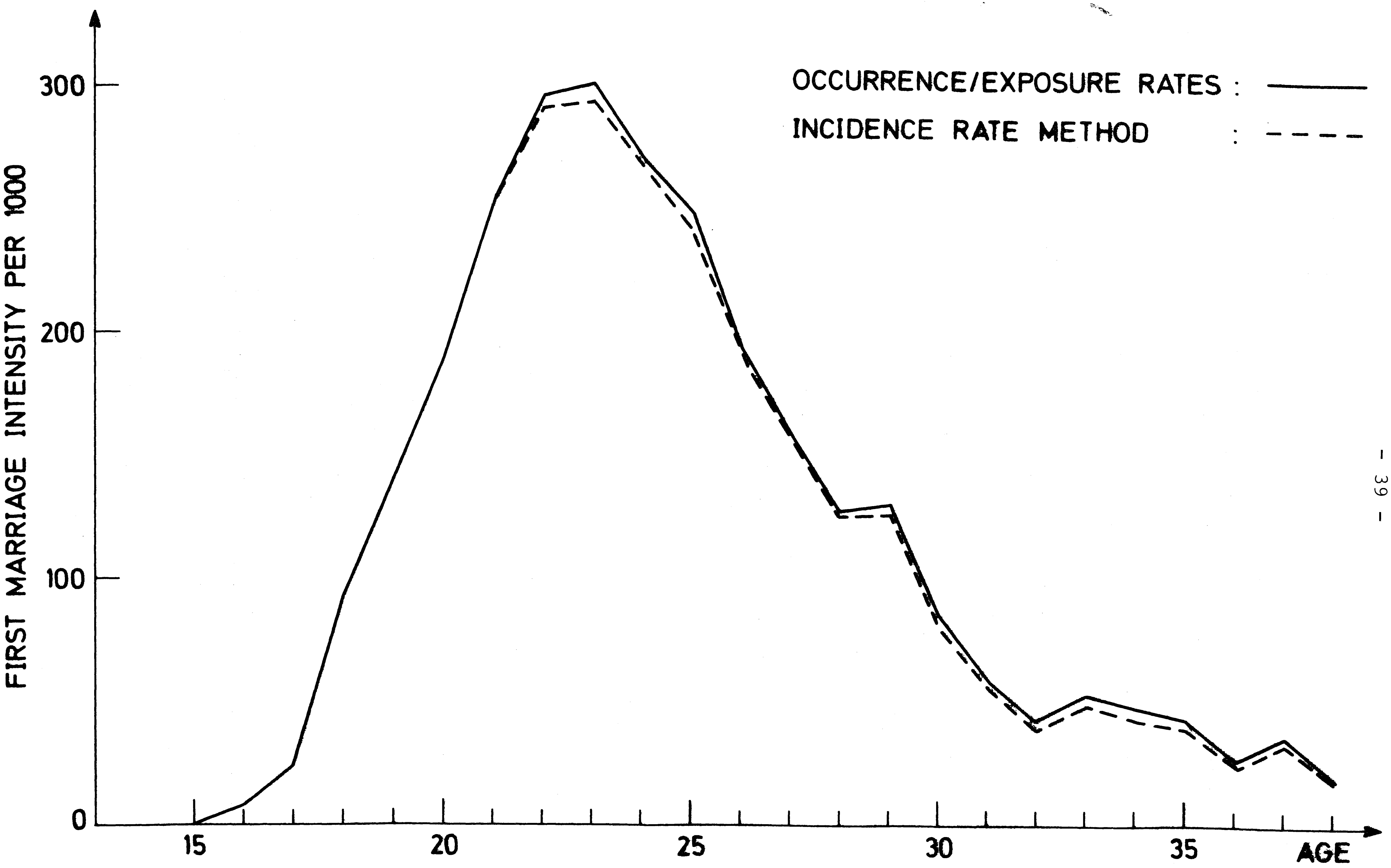

Fig. 2. Estimated first marriage intensity for women born in 1940 in Denmark.

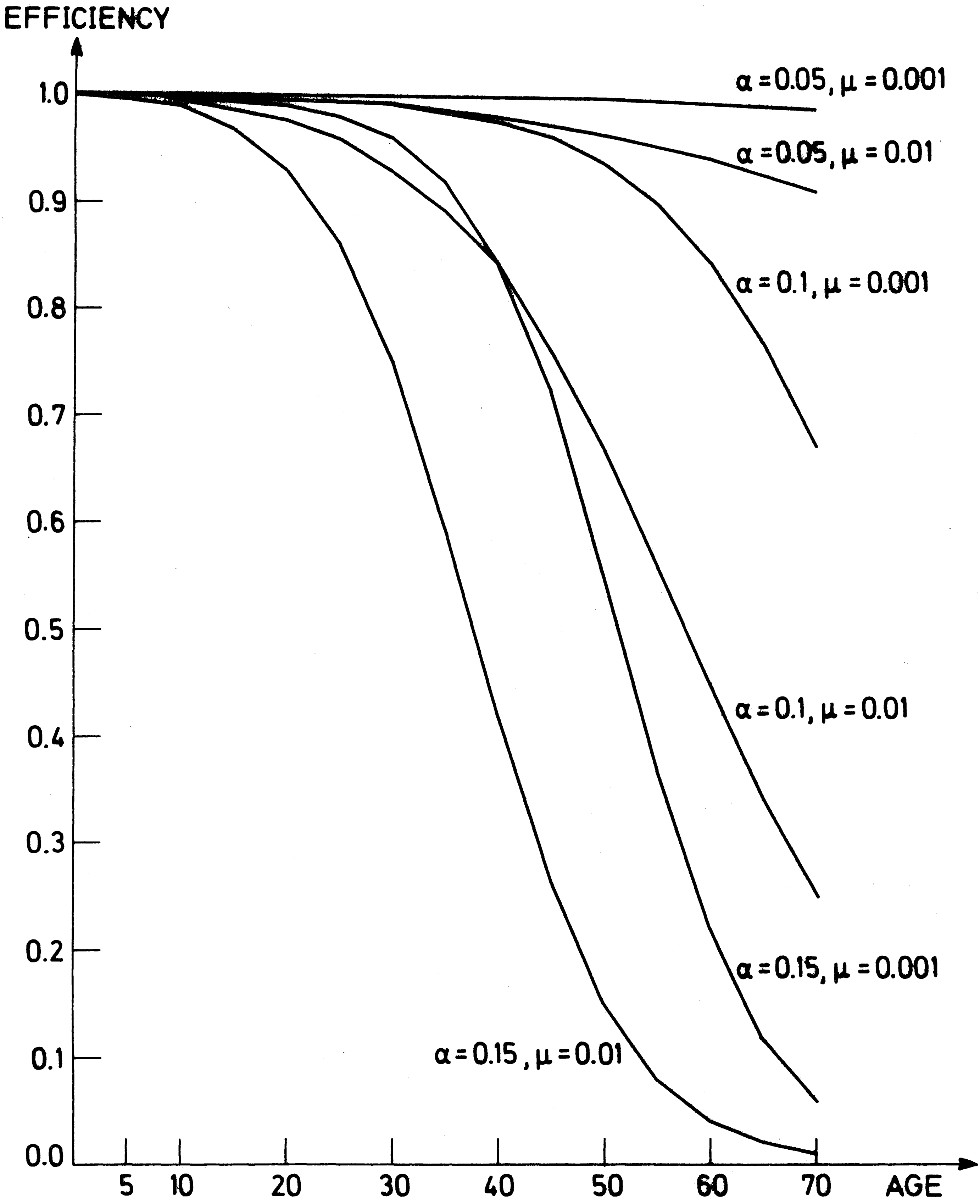

Fig. 3. Asymptotic relative efficiencies of the incidence rate method for the first marriage model assuming constant intensities throughout all ages.